Ph.D. Final Synopsis

Title: "Studies on Genetic Diversity of Koraga Populations of South West Coast of India – A Molecular Genetic Approach"

Reg. No. MU/EXB/Ph.D./CR.84/A.Z./2013-14/E.13 Dated 25.04.2014

Submitted by:

Research Scholar:
Vinuthalakshmi.K
Department of Applied Zoology
Mangalore University
Mangalagangothri

Research Supervisor:
Dr. Mohammed S. Mustak
Associate Professor
DepartmentofApplied Zoology
Mangalore University
Mangalagangothri

Place of Investigation:

Department of Applied Zoology, Mangalore University,

Mangalagangothri, Mangalore-574199

December -2019

Synopsis of the thesis submitted for the award of Ph.D. degree in Applied Zoology of Mangalore University, Mangalagangothri, Karnataka, India.

Title of Thesis

STUDIES ON GENETIC DIVERSITY OF KORAGA POPULATIONS OF SOUTH WEST COAST OF INDIA - A MOLECULAR GENETIC APPROACH

Candidate: **Vinutha Lakshmi.K**

India harbors vast human diversityconsisting of more than 4,635 anthropologically well-defined populations (castes, tribes and religious groups) of which 532 tribal communities who are the 'aboriginal' inhabitants of the sub-continent. Each population differs in terms of endogamy, language, culture, physical features, geographic and climatic position and genetic architecture. Evidence of ancient human dispersals and settlements in India has been extensively analysed in conventional and genomic analyses. A'tribe' is a social division in a traditional society consisting of families linked by social, economic, religious, or blood ties, with a common culture and dialect. A tribe possesses certain qualities and characteristics that make it a unique cultural, social, and political entity, constitute about 8.6% of the total population of India.These tribes speak about ~750 dialects that belongs to either one of the main four language groups, Austro-Asiatic(AA), Indo-Europeans(IE), Dravidians(DR) and Tibeto-Burmans(TB).Tribes of South India showed a reduction in genetic diversity and comprised large distances within a group as well as compared to other populations. This large genetic distance implies that tribes show less breeding probability with the neighboring populations and more isolated from the rest of the populations.The total number of tribal groups is estimated to be around 532, out of which 75 tribes are considered as primitive tribal groups. In India primitive tribes are characterized by geographically isolated and relatively cut-off from mainstream culture and exhibit backwardness in-terms of socio-economic and educational facilities.

The present study focused on the genetic origin, taste perception and alcoholic dependence of Koraga population -one of the primitive tribal groups residing in the South

west coast of India.They are small numbered, isolated populations practicing strict endogamy. Koragas are, a very quiet and inoffensive race, small and slight, the men seldom exceeding five feet six inches, black-skinned, thick-lipped, broad noses and flat, high-check bones and sloping foreheads, and with bushy rough hairs. Koragas have unique lifestyle habits compared to other tribal population, alcoholism and tobacco chewing are found to be common among them. They earn their livelihood through basket making. Based on hierarchy in Koraga tribes, they havebeen classified into different clans namely Onti, Kuntu and Soppu Koragas.

Koraga population is an unique and isolated population, most suitable for genetic model genetists to study the population with relevance to fouder events. To the best of our knowledge, no detailed research has been carried out in-detail concerning their genetic structure with respect to their ancestral origin,, taste perception and alcohol dependence. Therefore, the current study is focused on the following objectives to understand the genetic origin, genetic variation in taste perception and alcohol dependence of Koraga tribes.The significance findings are discussed in each chapter. Thepresent study to provide a new dimension in understanding the genetic structure of Koraga population.

The objectives of the present investigations are,

1. ***To trace the maternal ancestry of Koraga tribal population using Mitochondrial DNA as a genetic marker.***
2. ***To determine the bitter taste perception and associationof genetic variation of TAS2R38 gene polymorphism with lifestyle habits among the Koraga tribal population.***
3. ***To identify the genetic variation in ADH, ALDH and EPHX1genes and their association with alcoholism among Koraga tribal population.***

The research work is carried towards achieving these objectives that make the subject matter of the thesis. Therefore, the entire thesis is divided into five chapters.

Chapter 1. General Introduction

This chapter begins with a general account of the population of India, its richness in terms of linguistic, cultural, genetic diversity, followed by racial classification of tribal population of India. The chapter continues further by an overview of tribes; current

literature have been discussed in detail.The chapter also illustrates indetail about caste profile of Koragas. This chapter highlights the need of genetic analysis among Koraga tribes, also scope of the current study.

Chapter 2. Tracing the Maternal ancestry of Koraga tribal population by using mtDNA as Genetic marker

Mitochondrial DNA(mtDNA) has been a widely used tool in human evolutionary and population genetic studies over the past three decades for tracing maternal ancestry. Maternally inherited mtDNA with a lack of recombination, which retains notable genetic changes accumulated during the process of human evolution, human dispersion and human expansion. The studies showed that, mtDNA analysis is the best tool for understanding ancestral maternal lineage. Heterogeneity observed in the major mitochondrial DNA haplogroup (Hg), population differentiation estimates (*Fst)* and haplogroup frequencies are informative in finding founder events in the contemporary human population (Cavalli-Sforza et al., 2003). It is well established that different ethnic or population groups differ significantly in their mtDNA distributions and obtained data would be useful for elucidating the history of human population expansion across the world. In order to obtain maternal genetic prehistory and founder population of Koraga tribal population, isolated DNA were sequenced for the hypervariable region including HVR-1, HVR-II (16024-576) and partial coding region(14856-16023) for individuals belonging to different clans of Koraga tribal population (Soppu, Kuntu, and Onti clans).

The results highlighted the mtDNA pool of Koragas consists of total of 11 haplogroups , M3a2a, M3a1, U1a, U1a1a, U1a1a3,U2a1, M, M30, M30c, M40, M40a. These haplogroups are nested in the lineages of the macrohaplogroup M, N, R and U. The U1a and M3a1 were most common haplogroup among Koragas. The majority of Koraga population fell into the sub-lineages of the haplogroup 'U', specifically into U1a and U1a1a, U1a1a3, and U2a1, belongs to West-Eurasian specific lineages which accounts ~51%. In the current study, 'U' Haplogroup subclades having deep branches which are autochthonous to Koraga tribes of South West Coast of India. Further, study also reported with the presence of Indian specific mtDNA lineages, which consisting of haplogroup 'M' and its sub clades - M3a1, M3a2a,M30, M30c, M40,M40a harbors ~49% among Koragas.

The novelty of current study is that, for the first-time identified two major founder haplogroups among Koragas which harbored ~92% of their maternal ancestry. The major two haplogroups are U1a1a(specific to west Eurasia) M3a2a (origin of South Asia specific to India;) found among Koraga population. The Indian specific haplogroup M3a is more confined to South India.

Koraga tribal population showed high genetic diversity with a value of 0.840+/-0.0375. The distribution of Indian-specific haplogroups is irrespective of the language, caste and geography. The principal component analysis (PCA) and Fst results showed the Koragas population is isolated compared to the rest of other population, this may be attributed to the endogamy practices among the Koragas.

Previous studies have showed that, due to the genetic drift, there will be the possibility of high differentiation within the population with reduced diversity. However, according to previous studies, Koraga population showed more reduced diversity with the large genetic distances due to the genetic drift. Further, studies showed that inflow of genes during earlier period expansion of population rapidly resulted in high heterozygosity. The division of tribal population into small isolated, endogamous groups maintaining high genetic differentiation between the population groups and low diversity within the groups. the quite similar events can be seen in the Koraga population (Cordaux et al.,2003; Vishwanathan et al.,2004). The main novelty of the current study are:-*1. Two main founders of Koraga 2. Highly drifted population 3. First tribal population which has very high frequency of West Eurasian lineages.*

Chapter 3. Bitter taste perception and association of genetic variation of the TAS2R38 gene with alcoholism and Tobacco chewing

Bitter tasting ability may influence the choice of food and nutrient status of the individual. It is an essential evolutionary adaptation of the dietary behavior among the different human races for the avoidance of bitter, toxic compounds which are present in the form of food. Bitter taste sensitivity of Phenylthiocarbamide (PTC) or PROP (6-n Propylthiouracil) the sulfur-containing organic compound used as a genetic marker to trace the pattern of bitter taste response in a population. PTC tasting ability is genetically determined bimodal trait and classic example for a Mendelian recessive pattern of inheritance. TAS2R38

gene codes for bitter tasting ability, 3 SNPs are present at different nucleotide positions in the gene responsible for the variation to the dietary response. Presence of three functional single nucleotide polymorphisms (SNPs) in the *TAS2R38* gene at nucleotide position 145 (rs713598;C/G), 785(rs1726866;C/T), and 886 (rs10246939;G/A) encode for non-synonymous amino acid substitution at position 49 (Alanine/Proline as A49P), 262 (Valine/Alanine as V262A), and 296 (Isoleucine/Valine asI296V) responsible for changes in the responces for taste perception.

To determine the PTC taste sensitivity among the Koraga population, 0.025% PTC solution was used according to the modified method of Harris and Kalmus method(1949). The probability of percentage of non-tasters were found among higher age group. Pedigree analysis showed different patterns of inheritance comprising, recessive homozygous, dominant homozygous and heterozygous pattern of inheritance among (n=30) Koraga families. Other than the age factor, life style habits like alcoholism and tobacco chewing habits of Koragas were also affected taste sensitivity. Study recorded direct co-relationship between alcoholics with non-tasting ability among the Koragas with significant p-value (p-value $\leq$0.0000743).

Genotyping of all three variants of the TAS2R38 gene, gained minor allele frequencies (MAF) of 0.280, 0.373 and 0.424 at the position of G145C, T785C, and A886G respectively with an average mean value of 0.35 among the study population. Intriguingly, a world-wide population, minor allelic frequency (MAF) was ranging from 0.41-0.47 (Risso et al., 2016; Choi et al., 2016). The haplotype-based association analysis, showed that the major haplotype the PAV and AVI were present in the combination of PAV/PAV, PAV/AVI & AVI/AVI among the study population ~93% and the other rare combination of the haplotypes were observed ~ 7% population. The non-taster (AVI/AVI) haplotype was most commonly observed with a frequency of 2.6 times higher than that of homozygous taster PAV/PAV diplotype. Further, there was a positive association between non-taster haplotype AVI (71.4%) with alcoholic dependents(p-value$\leq$7.76061E^{-08}) and among the tobacco-chewers (53.0%). In conclusion, the analysis of genetic variants in the *TAS2R38* gene and haplotypes association showed that the high frequency of non-tasters might have influenced higher risk for alcohol dependence and tobacco chewing among Koraga tribal population.

Chapter 4. Determination of the genetic variation in ADH, ALDH and EPHX1 gene and their associations with alcoholism

The effect of alcohol induced organ damage depends upon the quantity of ethanol consumed and also depends on how effectively alcohol metabolism takes place with the help of body organs with the aid of enzymes. Major quantity of alcohol is catalyzed by the key enzymes called Alcohol dehydrogenase (ADH) and Aldehyde dehydrogenase (ALDH).In liver 80% alcohol degradation carried through oxidation in the liver. Ethanol metabolism involves various steps, which involves oxidation reaction, where ethanol is initially oxidized into acetaldehyde with aid of Alcohol dehydrogenases enzymes(ADHs) and it is rate limiting step of reaction in the cytosol. Acetaldehyde is further converted to acetate by Aldehyde dehydrogenase enzyme (ALDH), finally to water and carbon dioxide through the citric acid cycle. Thus, ADH and ALDH are the key enzymes involved in ethanol metabolism. The allelic variants of the genes coding for these enzymes exhibits polymorphism.Variations in the alleles of genes that coding for the alcohol metabolizing enzyme will affect the rate of metabolism of alcohol(Edenberg,2007). The distribution of the polymorphic alleles varies with ethnicity. Isomeric forms of alleles either protective or harmful to alcoholism. Due to the difference in distribution of the frequencies of Isoenzymes across the global ethnic races, the genes encoding for these enzymes acts as candidate genes to determine the variation in the response to drinking behaviour, to determine the development of alcoholism and alcohol related disorders.

Earlier study has showed that, extent of damage in hepatocytes of liver due to high alcohol consumption was comparatively lower among the Koragas when compared to general alcoholics(Praksh et al.,2010). However, there no report on molecular mechanisms behind the protective role to organ damage among the Koragas. Thus, the present study has aimed to identify the allelic and genotypic frequency of ADH1B(Arg47His;Exon3) and (Arg369Cys;Exon9),ADH1C(exon8), ALDH2(Exon12), EPHX1(Exon3,Exon4) genes and their association with alcoholism

Out of six markers, two marker coding for ADHIB gene (Arg47His;Exon3, Arg369Cys;Exon9) and one marker coding for ALDH2 gene (Glu487Lys;Exon12) not showed any polymorphic variation, remained in the wild homozygous condition as ADH1B*1 and ALDH2*1. On the

other hand, the remaining three markers coding for ADH1C(Ile350Val;Exon8) and EPHX1(Tyr113His;Exon3)and EPHX1(His139Arg;Exon4) genes exhibited polymorphism in both alcoholic cases and control samples. SNP(rs698) coding for ADH1C gene exhibited in all forms of alleles i.e. ADH1C*1(AA)' ADH1C*1/2'(AG)' and ADH1C*2(GG)' and similarly with SNPs(rs1051740 and rs2234922) of EPHX1 gene at both sites displayed all possibilities of genotypes combination, in EPHX1(Exon3) 'CC', 'CT' and 'TT' and EPHX1(Exon4) 'AA', 'AG' and 'GG' genotypes respectively.

Allelic and genotypic frequencies of ADH1C gene were exhibited differences interms of distribution among alcohol dependents(cases) of Koragas and non-alcoholics(controls). The genotypic frequencies of Ile(A/A) and Valine(G/G) in cases and controls were found to be 0.14 and 0.46 and 0.45 and 0.11 respectively. Allelic frequency of 'A' was recorded at low frequency among cases (0.3450) than control samples (0.6750), the likelihood of γ1(A/A)among controls was found to be 1.95 times higher than cases. Intrestingly, the allelic frequency of 'G'was recorded at high(0.6550) frequency among the cases of Koragas than the control group(0.3550), the likelihood of γ2 in cases was 2.01 times higher than the non-alcoholics(controls).

Various studies on ADH1C polymorphism has showed that alcohol-induced organ damage was associated with γ1γ1(AA genotype of ADH1C gene) and alcoholism was associated with γ2γ2(GG) of ADH1C gene marker. Further, studies on alcoholic caes among different populations showed, the presence of higher number of 'AA' genotype compared to 'GG' genotype. Contrasting to the previous findings, present study reports the presence of high frequency of 'GG' genotype(γ2γ2) among the alcoholics of Koraga population compared to non-alcoholics(controls). The presence of 'GG' in Koragas might help in protection for alcohol-induced organ damage and associated risk to alcoholism. This is the first report, to observe the presence of higher frequency of 'GG' genotypes in ADH1C gene among the alcohol dependents of Koragas instead of 'AA' genotypes contradictory to the earlier study findings.

In case of EPHX1 genes, Exon3 and Exon4 exhibited polymorphism at two sites. Allelic frequency and genotypic frequency compared between the groups at two sites, in Exon3(Tyr113 His), there were no significant differences in terms of allelic variation and

genotypic variation among alcoholics and non-alcoholics. However, at Exon4(His139Arg), there were differences in terms of the distribution of alleles among study groups(p-value is 0.0125). The allelic frequencies of Histidine(A) and Arginine(G) were 0.3511and 0.5125 and 0.6489 and 0.4875 among alcoholics and non-alcoholics respectively. Thus, EPHX1 gene at both sites showed an independent effect among study groups. EPHX1 activity is classified as Low(Tyr113His/ His139His or His113His/His139Arg orHis113His/His139His), Intermediate(Tyr113Tyr/His139Hisor Tyr113His/His139Arg or His113His/Arg139Arg), High(Tyr113Tyr/His139Arg or Tyr113Tyr/Arg139Arg or Tyr113His/Arg139Arg. The present study observed, the high activity samples of EPHX1 gene were prevalent among the Koragas compared to the control(non-alcoholics) population, genotypes combinations of high activity samples will lead to efficient detoxification of carcinogens which created during alcohol metabolism.

Further, to understand the genetic risk factors associated with alcoholism among the Koraga tribal populations performed the genetic association analysis. Out of 6 mutations, identified 'rs698' as genetic risk factor for alcoholism (p-value: allelic = 5.05×10^{-9} and genotypic = 2.34×10^{-7}). Likelihood of being alcoholic in individual with "G", "GG" and "GA". Thus, the current study showed 'ADH1C' gene (rs698) as a candidate gene to predict genetic risk factor associated with alcoholism(p-value: allelic = 5.05×10-9 and genotypic = 2.34×10-7) among Koraga population.

Chapter 5 : The thesis ends with a comprehensive **Summary and Conclusions** with future prespetives.

In a nut shell the study provides following significant findings.

1. **The present study has generated comprehensive data on maternal ancestry of Koragas using mtDNA marker to trace maternal ancestry. mtDNA pool of Koragas consists of several deep rooted autochotonous lineages of macrohaplogroup 'M' and 'U' suggesting *in-situ* origin of mito haplogroups in Koragas. MtDNA haplogroup analysis among Koragas revealed that Koragas maternal root consists of higher frequencies of West Eurasian specific lineages like U1a1a , U1a1a3, U2a1 and this type of specific lineages recorded only in Koraga tribal population of South India and it is unique to these tribes. Other than the West Eurasian specificity, Indian**

specific mito lineages such as M3a1a, M3a2a of South Asian origin and Indian specific lineages were also identified among Koragas. Thus, Koragas maternal line might have admixed with West Eurasian and Indian specific lineages. Further, present study evidenced similar genetic structure among the three clans of Koragas(Chapter 2) .

2. **Phenotypic, genotypic and haplotypic association analysis revealed the presence of high frequency of non-tasters among Koraga tribal population. The non-taster haplotype AVI is more prevalent than taster haplotype PAV among the alcohol dependents and tobacco chewers suggesting, Koragas have higher risk to alcohol consumption and tobacco chewing habits (Chapter 3).**

3. **Out of six markers, SNP marker (rs698) in ADH1C gene showed the association with alcoholic dependence in Koragas and further, observed ADH1C*2 (GG) at high frequency which is a protective allele against alcohol induced organ damage, Further, polymorphism in EPHX1 gene showed the prevalence of high activity of EPHX1 genes among alcohol dependents of Koragas (Chapter 4).**

In the thesis the significance of the results are discussed in detail.

Paper Published

Vinuthalakshmi K, Sheik N., Mustak M.S.,2019.*TAS2R38* gene polymorphism and its association with taste perception, alcoholism and tobacco chewing among the Koraga -a primitive tribalpopulation of Southwest coast of India. *Meta Gene*, 20,1-8.

Paper Presentations:

Vinutha Lakshmi.K, R Rajkumar, M Naveen, K Thangaraj, M S Mustak, presented a paper on**"ADH1 and ALDH2 gene polymorphisms among the Koraga populationof South Karnataka, India"**. International conference on Human Genetics and 40th annual meeting on Indian society of Human genetics (ISHG-2015), Mumbai, India, 27th- 30th January 2015.

Vinutha Lakshmi.K, M S Mustak, presented a paper on**"Genetic analysis of Phenylthiocarbamide taste perception among Koraga tribal population of Dakshina Kannada, Udupi districts of Karnataka".** National level conference on 'Recent advances on Animal Sciences', Department of Applied Zoology, Mangalore University, 13th - 15th February 2015.

K Vinuthalakshmi, S Nizamuddin, M Naveen, K Thangaraj, M S Mustak, presented a paper on**"Association study of the Ile349Val polymorphism of the gene *ADH1C* and alcoholism of Tribal population of Karnataka.** National Conference on "Frontiers in Genetics and Genomics", Department of Genomic Science, Central University of Kerala, Kasaragod, Kerala on 07-08th April 2016.

K Vinuthalakshmi, Nagarjuna P, M S Mustak, presented a paper on**"Bitter taste perception of Phenylthiocarbamide among the Koraga population through the analysis of TAS2R38 gene.** National Conference on "Reaching theunreached through Science and Technology", Department of Applied Zoology, Mangalore University, Mangalagangothri, Mangaluru on 08-09th September 2017.

Research Scholar
Vinutha Lakshmi.K

Research Guide
Dr. M. S. Mustak